

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



T1590

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/12, 15/11, 15/85, 5/10, C07K 14/705		A1	(11) International Publication Number: WO 99/37765
			(43) International Publication Date: 29 July 1999 (29.07.99)
(21) International Application Number: PCT/EP99/00420		(81) Designated States: JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(22) International Filing Date: 25 January 1999 (25.01.99)		Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(30) Priority Data: 98300549.7 27 January 1998 (27.01.98) EP 9823421.4 26 October 1998 (26.10.98) GB 9901209.8 20 January 1999 (20.01.99) GB			
(71) Applicant: SMITHKLINE BEECHAM PLC [GB/GB]; New Horizons Court, Brentford, Middlesex TW8 9EP (GB).			
(72) Inventors: DAVIS, John, Beresford; SmithKline Beecham Pharmaceuticals, New Frontiers Science Park South, Third Avenue, Harlow, Essex CM19 5AW (GB). DUCKWORTH, David, Malcolm; SmithKline Beecham Pharmaceuticals, New Frontiers Science Park South, Third Avenue, Harlow, Essex CM19 5AW (GB). HAYES, Philip, David; SmithKline Beecham Pharmaceuticals, New Frontiers Science Park South, Third Avenue, Harlow, Essex CM19 5AW (GB).			
(74) Agent: CONNELL, Anthony, Christopher; SmithKline Beecham, Corporate Intellectual Property, Two New Horizons Court, Brentford, Middlesex TW8 9EP (GB).			

(54) Title: HUMAN VANILLOID RECEPTOR HOMOLOGUES

(57) Abstract

VANILREP2 polypeptides and polynucleotides and methods for producing such polypeptides by recombinant techniques are disclosed. Also disclosed are methods for utilizing VANILREP2 polypeptides and polynucleotides in therapy, and diagnostic assays for such.

BEST AVAILABLE COPY

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		

HUMAN VANILLOID RECEPTOR HOMOLOGUES

Field of the Invention

This invention relates to newly identified polypeptides and polynucleotides encoding such polypeptides, to their use in therapy and in identifying compounds which may be agonists, 5 antagonists and /or inhibitors which are potentially useful in therapy, and to production of such polypeptides and polynucleotides.

Background of the Invention

The drug discovery process is currently undergoing a fundamental revolution as it embraces 10 'functional genomics', that is, high throughput genome- or gene-based biology. This approach as a means to identify genes and gene products as therapeutic targets is rapidly superceding earlier approaches based on 'positional cloning'. A phenotype, that is a biological function or genetic disease, would be identified and this would then be tracked back to the responsible gene, based on its genetic map position.

15 Functional genomics relies heavily on high-throughput DNA sequencing technologies and the various tools of bioinformatics to identify gene sequences of potential interest from the many molecular biology databases now available. There is a continuing need to identify and characterise further genes and their related polypeptides/proteins, as targets for drug discovery.

20 Summary of the Invention

The present invention relates to VANILREP2, in particular VANILREP2 polypeptides and VANILREP2 polynucleotides, recombinant materials and methods for their production. In another aspect, the invention relates to methods for using such polypeptides and polynucleotides, including the treatment of pain, chronic pain, neuropathic pain, postoperative pain, rheumatoid 25 arthritic pain, neuralgia, neuropathies, algesia, nerve injury, ischaemia, neurodegeneration, stroke, incontinence and inflammatory disorders, hereinafter referred to as "the Diseases", amongst others. In a further aspect, the invention relates to methods for identifying agonists and antagonists/inhibitors using the materials provided by the invention, and treating conditions associated with VANILREP2 imbalance with the identified compounds. In a still further aspect, 30 the invention relates to diagnostic assays for detecting diseases associated with inappropriate VANILREP2 activity or levels.

Description of the Invention

In a first aspect, the present invention relates to VANILREP2 polypeptides. These include the polypeptide of SEQ ID NO:2 and polymorphic variants thereof, for example PVP-1, the polypeptide of SEQ ID NO:6. Such peptides include isolated polypeptides comprising an amino acid sequence which has at least 95% identity, preferably at least 97-99% identity, to that of SEQ ID NO:2 or SEQ ID NO:6 over the entire length of SEQ ID NO:2 or SEQ ID NO:6 respectively. Such polypeptides include those comprising the amino acid of SEQ ID NO:2 or SEQ ID NO:6.

5 Further peptides of the present invention include isolated polypeptides in which the amino acid sequence has at least 95% identity, preferably at least 97-99% identity, to the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:6 over the entire length of SEQ ID NO:2 or SEQ ID NO:6. Such polypeptides include the polypeptide of SEQ ID NO:2 or SEQ ID NO:6.

10 Further peptides of the present invention include isolated polypeptides encoded by a polynucleotide comprising the sequence contained in SEQ ID NO:1 or SEQ ID NO:5.

15 Polypeptides of the present invention are believed to be members of the ion channel family of polypeptides. They are therefore of interest because they are associated with the mechanism of action of capsaicin (a vanilloid compound), a constituent of chilli peppers. Capsaicin elicits a sensation of burning pain by selectively activating sensory neurons that convey information about noxious stimuli to the central nervous system. The channels are permeable to cations and exhibit a notable preference for divalent cations, particularly calcium ions. The level of calcium ion 20 permeability exceeds that observed for most non-selective cation channels and is similar to values observed for NMDA-type glutamate receptors and alpha7 nicotinic acetylcholine receptors, both of which are noted for this property. These properties are hereinafter referred to as "VANILREP2 activity" or "VANILREP2 polypeptide activity" or "biological activity of VANILREP2". Also included amongst these activities are antigenic and immunogenic activities of said VANILREP2 25 polypeptides, in particular the antigenic and immunogenic activities of the polypeptides of SEQ ID NO:2 or SEQ ID NO:6. Preferably, a polypeptide of the present invention exhibits at least one biological activity of VANILREP2.

30 The polypeptides of the present invention may be in the form of the "mature" protein or may be a part of a larger protein such as a precursor or a fusion protein. It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification such as multiple histidine residues, or an additional sequence for stability during recombinant production.

The present invention also includes variants of the aforementioned polypeptides, that is polypeptides that vary from the referents by conservative amino acid substitutions, whereby a

residue is substituted by another with like characteristics. Typical such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; and among the basic residues Lys and Arg; or aromatic residues Phe and Tyr. Particularly preferred are variants in which several, 5-10, 1-5, 1-3, 1-2 or 1 amino acids are substituted, deleted, or added 5 in any combination.

Polypeptides of the present invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

10 In a further aspect, the present invention relates to VANILREP2 polynucleotides. Such polynucleotides include isolated polynucleotides comprising a nucleotide sequence encoding a polypeptide which has at least 95% identity to the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:6, over the entire length of SEQ ID NO:2 or SEQ ID NO:6, respectively. In this regard, polypeptides which have at least 97% identity are highly preferred, whilst those with at least 98-99% 15 identity are more highly preferred, and those with at least 99% identity are most highly preferred. Such polynucleotides include a polynucleotide comprising the nucleotide sequence contained in SEQ ID NO:1 or SEQ ID NO:5 encoding the polypeptide of SEQ ID NO:2 or SEQ ID NO:6, respectively.

20 Further polynucleotides of the present invention include isolated polynucleotides comprising a nucleotide sequence that has at least 95% identity to a nucleotide sequence encoding a polypeptide of SEQ ID NO:2 or SEQ ID NO:6, over the entire coding region. In this regard, polynucleotides which have at least 97% identity are highly preferred, whilst those with at least 98-99% identity are more highly preferred, and those with at least 99% identity are most highly preferred.

25 Further polynucleotides of the present invention include isolated polynucleotides comprising a nucleotide sequence which has at least 95% identity to SEQ ID NO:1 or SEQ ID NO:5 over the entire length of SEQ ID NO:1 or SEQ ID NO:5, respectively. In this regard, polynucleotides which have at least 97% identity are highly preferred, whilst those with at least 98-99% identiy are more highly preferred, and those with at least 99% identity are most highly 30 preferred. Such polynucleotides include a polynucleotide comprising the polynucleotide of SEQ ID NO:1 or SEQ ID NO:5 as well as the polynucleotide of SEQ ID NO:1 or SEQ ID NO:5, respectively.

The invention also provides polynucleotides which are complementary to all the above described polynucleotides.

The nucleotide sequences of SEQ ID NO:1 or SEQ ID NO:5 show homology with rat vanilloid receptor VR1(M. J. Caterina et al., *Nature* 389: 816-824, 1997). The nucleotide sequence of SEQ ID NO:1 is a cDNA sequence and comprises a polypeptide encoding sequence (nucleotide 5 to 2296) encoding a polypeptide of 764 amino acids, the polypeptide of SEQ ID NO:2. The 5 nucleotide sequence encoding the polypeptide of SEQ ID NO:2 may be identical to the polypeptide encoding sequence contained in SEQ ID NO:1 or it may be a sequence other than the one contained in SEQ ID NO:1, which, as a result of the redundancy (degeneracy) of the genetic code, also encodes the polypeptide of SEQ ID NO:2. The polypeptide of the SEQ ID NO:2 is structurally related to other proteins of the ion channel family, having homology and/or structural 10 similarity with rat vanilloid receptor VR1(M. J. Caterina et al., *Nature* 389: 816-824, 1997).

The nucleotide sequence of SEQ ID NO:5 is a cDNA sequence and comprises a polypeptide encoding sequence (nucleotide 5 to 2293) encoding a polypeptide of 763 amino acids, the polypeptide of SEQ ID NO:6. The nucleotide sequence encoding the polypeptide of SEQ ID NO:6 may be identical to the polypeptide encoding sequence contained in SEQ ID NO:5 or it 15 may be a sequence other than the one contained in SEQ ID NO:5, which, as a result of the redundancy (degeneracy) of the genetic code, also encodes the polypeptide of SEQ ID NO:6. The polypeptide of the SEQ ID NO:6 is structurally related to other proteins of the ion channel family, having homology and/or structural similarity with rat vanilloid receptor VR1(M. J. Caterina et al., *Nature* 389: 816-824, 1997).

20 Preferred polypeptides and polynucleotides of the present invention are expected to have, *inter alia*, similar biological functions/properties to their homologous polypeptides and polynucleotides. Furthermore, preferred polypeptides and polynucleotides of the present invention have at least one VANILREP2 activity.

The present invention also relates to partial or other polynucleotide and polypeptide 25 sequences which were first identified prior to the determination of the corresponding full length sequences of SEQ ID NO:1, SEQ ID NO:5, SEQ ID NO:2 and SEQ ID NO:6.

Accordingly, in a further aspect, the present invention provides for an isolated polynucleotide which:

- (a) comprises a nucleotide sequence which has at least 95% identity, preferably at least 97-99% 30 identity to SEQ ID NO:3 over the entire length of SEQ ID NO:3;
- (b) a nucleotide sequence which has at least 95% identity preferably at least 97-99% identity, to SEQ ID NO:3 over the entire length of SEQ ID NO:3;
- (c) the polynucleotide of SEQ ID NO:3; or

(d) a nucleotide sequence encoding a polypeptide which has at least 95% identity preferably at least 97-99% identity, to the amino acid sequence of SEQ ID NO:4, over the entire length of SEQ ID NO:4;

as well as the polynucleotide of SEQ ID NO:3.

5 The present invention further provides for a polypeptide which:

(a) comprises an amino acid sequence which has at least 95% identity, preferably at least 97-99% identity, to that of SEQ ID NO:4 over the entire length of SEQ ID NO:4;

(b) has an amino acid sequence which is at least 95% identity, preferably at least 97-99% identity, to the amino acid sequence of SEQ ID NO:4 over the entire length of SEQ ID NO:4;

10 (c) comprises the amino acid of SEQ ID NO:4; and

(d) is the polypeptide of SEQ ID NO:4;

as well as polypeptides encoded by a polynucleotide comprising the sequence contained in SEQ ID NO:3.

15 The nucleotide sequence of SEQ ID NO:3 and the peptide sequence encoded thereby are derived from EST (Expressed Sequence Tag) sequences. It is recognised by those skilled in the art that there will inevitably be some nucleotide sequence reading errors in EST sequences (see Adams, M.D. *et al.*, *Nature* 377 (supp) 3, 1995). Accordingly, the nucleotide sequence of SEQ ID NO:3 and the peptide sequence encoded therefrom are therefore subject to the same inherent limitations in sequence accuracy.

20 Polynucleotides of the present invention may be obtained, using standard cloning and screening techniques, from a cDNA library derived from mRNA in cells of human placenta, brain, heart, melanocyte, liver, spleen, foreskin, lung, parathyroid, tonsil, whole embryo, Jurkat T-cells, B-cells and testes (for example Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)). Polynucleotides of the 25 invention can also be obtained from natural sources such as genomic DNA libraries or can be synthesized using well known and commercially available techniques.

When polynucleotides of the present invention are used for the recombinant production of polypeptides of the present invention, the polynucleotide may include the coding sequence for the mature polypeptide, by itself; or the coding sequence for the mature polypeptide in reading 30 frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence, or other fusion peptide portions. For example, a marker sequence which facilitates purification of the fused polypeptide can be encoded. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al.*, *Proc Natl Acad Sci USA*

(1989) 86:821-824, or is an HA tag. The polynucleotide may also contain non-coding 5' and 3' sequences, such as transcribed, non-translated sequences, splicing and polyadenylation signals, ribosome binding sites and sequences that stabilize mRNA.

Further embodiments of the present invention include polynucleotides encoding polypeptide variants which comprise the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:6, respectively and in which several, for instance from 5 to 10, 1 to 5, 1 to 3, 1 to 2 or 1, amino acid residues are substituted, deleted or added, in any combination.

Polynucleotides which are identical or sufficiently identical to a nucleotide sequence contained in SEQ ID NO:1 or SEQ ID NO:5, respectively, may be used as hybridization probes for cDNA and genomic DNA or as primers for a nucleic acid amplification (PCR) reaction, to isolate full-length cDNAs and genomic clones encoding polypeptides of the present invention and to isolate cDNA and genomic clones of other genes (including genes encoding paralogs from human sources and orthologs and paralogs from species other than human) that have a high sequence similarity to SEQ ID NO:1 or SEQ ID NO:5. Typically these nucleotide sequences are 70% identical, preferably 80% identical, more preferably 90% identical, most preferably 95% identical to that of the referent. The probes or primers will generally comprise at least 15 nucleotides, preferably, at least 30 nucleotides and may have at least 50 nucleotides. Particularly preferred probes will have between 30 and 50 nucleotides. Particularly preferred primers will have between 20 and 25 nucleotides.

A polynucleotide encoding a polypeptide of the present invention, including homologs from species other than human, may be obtained by a process which comprises the steps of screening an appropriate library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NO: 1 or SEQ ID NO:5 or a fragment thereof; and isolating full-length cDNA and genomic clones containing said polynucleotide sequence. Such hybridization techniques are well known to the skilled artisan. Preferred stringent hybridization conditions include overnight incubation at 42°C in a solution comprising: 50% formamide, 5xSSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10 % dextran sulfate, and 20 microgram/ml denatured, sheared salmon sperm DNA; followed by washing the filters in 0.1x SSC at about 65°C. Thus the present invention also includes polynucleotides obtainable by screening an appropriate library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NO:1 or SEQ ID NO:5, or a fragment thereof.

The skilled artisan will appreciate that, in many cases, an isolated cDNA sequence will be incomplete, in that the region coding for the polypeptide is short at the 5' end of the cDNA. This is a consequence of reverse transcriptase, an enzyme with inherently low 'processivity' (a measure of the ability of the enzyme to remain attached to the template during the polymerisation

reaction), failing to complete a DNA copy of the mRNA template during 1st strand cDNA synthesis.

There are several methods available and well known to those skilled in the art to obtain full-length cDNAs, or extend short cDNAs, for example those based on the method of Rapid

5 Amplification of cDNA ends (RACE) (see, for example, Frohman *et al.*, PNAS USA 85, 8998-9002, 1988). Recent modifications of the technique, exemplified by the Marathon™ technology (Clontech Laboratories Inc.) for example, have significantly simplified the search for longer cDNAs. In the Marathon™ technology, cDNAs have been prepared from mRNA extracted from a chosen tissue and an 'adaptor' sequence ligated onto each end. Nucleic acid amplification
10 (PCR) is then carried out to amplify the 'missing' 5' end of the cDNA using a combination of gene specific and adaptor specific oligonucleotide primers. The PCR reaction is then repeated using 'nested' primers, that is, primers designed to anneal within the amplified product (typically an adaptor specific primer that anneals further 3' in the adaptor sequence and a gene specific primer that anneals further 5' in the known gene sequence). The products of this reaction can then be
15 analysed by DNA sequencing and a full-length cDNA constructed either by joining the product directly to the existing cDNA to give a complete sequence, or carrying out a separate full-length PCR using the new sequence information for the design of the 5' primer.

Recombinant polypeptides of the present invention may be prepared by processes well known in the art from genetically engineered host cells comprising expression systems.

20 Accordingly, in a further aspect, the present invention relates to expression systems which comprise a polynucleotide or polynucleotides of the present invention, to host cells which are genetically engineered with such expression systems and to the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention.

25 For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof for polynucleotides of the present invention. Introduction of polynucleotides into host cells can be effected by methods described in many standard laboratory manuals, such as Davis *et al.*, *Basic Methods in Molecular Biology* (1986) and Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold
30 Spring Harbor, N.Y. (1989). Preferred such methods include, for instance, calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection.

Representative examples of appropriate hosts include bacterial cells, such as *Streptococci*, *Staphylococci*, *E. coli*, *Streptomyces* and *Bacillus subtilis* cells; fungal cells, such as yeast cells and *Aspergillus* cells; insect cells such as *Drosophila S2* and *Spodoptera Sf9* cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, HEK 293 and Bowes melanoma cells; and plant cells.

5 A great variety of expression systems can be used, for instance, chromosomal, episomal and virus-derived systems, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression systems may contain control regions that regulate as well as engender expression. Generally, any system or vector which is able to maintain, propagate or express a polynucleotide to produce a polypeptide in a host may be used. The appropriate nucleotide sequence may be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al.*, Molecular Cloning, A Laboratory Manual (supra). Appropriate secretion signals may be incorporated into the desired polypeptide to allow secretion of the translated protein into the lumen of the endoplasmic reticulum, the periplasmic space or the extracellular environment. These signals may be endogenous to the polypeptide or they may be heterologous signals.

10 15 20 25 30 If a polypeptide of the present invention is to be expressed for use in screening assays, it is generally preferred that the polypeptide be produced at the surface of the cell. In this event, the cells may be harvested prior to use in the screening assay. If the polypeptide is secreted into the medium, the medium can be recovered in order to recover and purify the polypeptide. If produced intracellularly, the cells must first be lysed before the polypeptide is recovered.

Polypeptides of the present invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during intracellular synthesis, isolation and or purification.

This invention also relates to the use of polynucleotides of the present invention as diagnostic reagents. Detection of a mutated form of the gene characterised by the polynucleotide of

SEQ ID NO:1 or SEQ ID NO:5, respectively, which is associated with a dysfunction will provide a diagnostic tool that can add to, or define, a diagnosis of a disease, or susceptibility to a disease, which results from under-expression, over-expression or altered spatial or temporal expression of the gene. Individuals carrying mutations in the gene may be detected at the DNA level by a variety 5 of techniques.

Nucleic acids for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR or other amplification techniques prior to analysis. RNA or cDNA may also be used in similar fashion. Deletions and insertions can be 10 detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to labeled VANILREP2 nucleotide sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence differences may also be detected by alterations in electrophoretic mobility of DNA fragments in gels, with or without 15 denaturing agents, or by direct DNA sequencing (ee, e.g., Myers *et al.*, *Science* (1985) 230:1242). Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (see Cotton *et al.*, *Proc Natl Acad Sci USA* (1985) 85: 4397-4401). In another embodiment, an array of oligonucleotides probes comprising VANILREP2 nucleotide sequence or fragments thereof can be constructed to conduct 20 efficient screening of e.g., genetic mutations. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability (see for example: M.Chee *et al.*, *Science*, Vol 274, pp 610-613 (1996)).

The diagnostic assays offer a process for diagnosing or determining a susceptibility to the 25 Diseases through detection of mutation in the VANILREP2 gene by the methods described. In addition, such diseases may be diagnosed by methods comprising determining from a sample derived from a subject an abnormally decreased or increased level of polypeptide or mRNA. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art for the quantitation of polynucleotides, such as, for example, nucleic acid 30 amplification, for instance PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods. Assay techniques that can be used to determine levels of a protein, such as a polypeptide of the present invention, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

Thus in another aspect, the present invention relates to a diagnostic kit which comprises:

- (a) a polynucleotide of the present invention, preferably the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO:5, respectively, or a fragment thereof;
- (b) a nucleotide sequence complementary to that of (a);
- 5 (c) a polypeptide of the present invention, preferably the polypeptide of SEQ ID NO:2 or SEQ ID NO:6, respectively or a fragment thereof; or
- (d) an antibody to a polypeptide of the present invention, preferably to the polypeptide of SEQ ID NO:2 or SEQ ID NO:6, respectively.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component. Such a kit will be of use in diagnosing a disease or susceptibility to a disease, particularly pain, chronic pain, neuropathic pain, postoperative pain, rheumatoid arthritic pain, neuralgia, neuropathies, algesia, nerve injury, ischaemia, neurodegeneration, stroke, incontinence and inflammatory disorders, amongst others.

15 The nucleotide sequences of the present invention are also valuable for chromosomal localisation. The sequence is specifically targeted to, and can hybridize with, a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important first step in correlating those sequences with gene associated disease. Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such 20 data are found in, for example, V. McKusick, Mendelian Inheritance in Man (available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

25 The differences in the cDNA or genomic sequence between affected and unaffected individuals can also be determined. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

30 The nucleotide sequences of the present invention are also valuable for tissue localisation. Such techniques allow the determination of expression patterns of the human VANILREP2 polypeptides in tissues by detection of the mRNAs that encode them. These techniques include in situ hybridization techniques and nucleotide amplification techniques, for example PCR. Such techniques are well known in the art. Results from these studies provide an indication of the normal functions of the polypeptides in the organism. In addition, comparative studies of the normal expression pattern of human VANILREP2 mRNAs with that of mRNAs

encoded by a human VANILREP2 gene provide valuable insights into the role of mutant human VANILREP2 polypeptides, or that of inappropriate expression of normal human VANILREP2 polypeptides, in disease. Such inappropriate expression may be of a temporal, spatial or simply quantitative nature.

5 The polypeptides of the invention or their fragments or analogs thereof, or cells expressing them, can also be used as immunogens to produce antibodies immunospecific for polypeptides of the present invention. The term "immunospecific" means that the antibodies have substantially greater affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art.

10 Antibodies generated against polypeptides of the present invention may be obtained by administering the polypeptides or epitope-bearing fragments, analogs or cells to an animal, preferably a non-human animal, using routine protocols. For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., *Nature* (1975) 256:495-15 497), the trioma technique, the human B-cell hybridoma technique (Kozboret *et al.*, *Immunology Today* (1983) 4:72) and the EBV-hybridoma technique (Cole *et al.*, *Monoclonal Antibodies and Cancer Therapy*, 77-96, Alan R. Liss, Inc., 1985).

20 Techniques for the production of single chain antibodies, such as those described in U.S. Patent No. 4,946,778, can also be adapted to produce single chain antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms, including other mammals, may be used to express humanized antibodies.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or to purify the polypeptides by affinity chromatography.

25 Antibodies against polypeptides of the present invention may also be employed to treat the Diseases, amongst others.

In a further aspect, the present invention relates to genetically engineered soluble fusion proteins comprising a polypeptide of the present invention, or a fragment thereof, and various portions of the constant regions of heavy or light chains of immunoglobulins of various subclasses (IgG, IgM, IgA, IgE). Preferred as an immunoglobulin is the constant part of the 30 heavy chain of human IgG, particularly IgG1, where fusion takes place at the hinge region. In a particular embodiment, the Fc part can be removed simply by incorporation of a cleavage sequence which can be cleaved with blood clotting factor Xa. Furthermore, this invention relates to processes for the preparation of these fusion proteins by genetic engineering, and to the use thereof for drug screening, diagnosis and therapy. A further aspect of the invention also relates to

polynucleotides encoding such fusion proteins. Examples of fusion protein technology can be found in International Patent Application Nos. WO94/29458 and WO94/22914.

Another aspect of the invention relates to a method for inducing an immunological response in a mammal which comprises inoculating the mammal with a polypeptide of the present invention, adequate to produce antibody and/or T cell immune response to protect said animal from the Diseases hereinbefore mentioned, amongst others. Yet another aspect of the invention relates to a method of inducing immunological response in a mammal which comprises, delivering a polypeptide of the present invention *via* a vector directing expression of the polynucleotide and coding for the polypeptide *in vivo* in order to induce such an immunological response to produce antibody to protect said animal from diseases.

A further aspect of the invention relates to an immunological/vaccine formulation (composition) which, when introduced into a mammalian host, induces an immunological response in that mammal to a polypeptide of the present invention wherein the composition comprises a polypeptide or polynucleotide of the present invention. The vaccine formulation may further comprise a suitable carrier. Since a polypeptide may be broken down in the stomach, it is preferably administered parenterally (for instance, subcutaneous, intramuscular, intravenous, or intradermal injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

Polypeptides of the present invention are responsible for one or more biological functions, including one or more disease states, in particular the Diseases hereinbefore mentioned. It is therefore desirous to devise screening methods to identify compounds which stimulate or which inhibit the function of the polypeptide. Accordingly, in a further aspect, the present invention provides for a method of screening compounds to identify those which stimulate or which inhibit the function of the polypeptide. In general, agonists or antagonists may be employed for therapeutic and prophylactic purposes for such Diseases as hereinbefore mentioned. Compounds may be identified from a variety of sources, for example, cells, cell-free preparations, chemical libraries, and

natural product mixtures. Such agonists, antagonists or inhibitors so-identified may be natural or modified substrates, ligands, receptors, enzymes, etc., as the case may be, of the polypeptide; or may be structural or functional mimetics thereof (see Coligan *et al.*, Current Protocols in Immunology 1(2):Chapter 5 (1991)).

5 The screening method may simply measure the binding of a candidate compound to the polypeptide, or to cells or membranes bearing the polypeptide, or a fusion protein thereof by means of a label directly or indirectly associated with the candidate compound. Alternatively, the screening method may involve competition with a labeled competitor. Further, these screening methods may test whether the candidate compound results in a signal generated by activation or
10 inhibition of the polypeptide, using detection systems appropriate to the cells bearing the polypeptide. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed. Constitutively active polypeptides may be employed in screening methods for inverse agonists or inhibitors, in the absence of an agonist or inhibitor, by testing whether the candidate
15 compound results in inhibition of activation of the polypeptide. Further, the screening methods may simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide of the present invention, to form a mixture, measuring VANILREP2 activity in the mixture, and comparing the VANILREP2 activity of the mixture to a standard. Fusion proteins, such as those made from Fc portion and VANILREP2 polypeptide, as hereinbefore described, can
20 also be used for high-throughput screening assays to identify antagonists for the polypeptide of the present invention (see D. Bennett *et al.*, J Mol Recognition, 8:52-58 (1995); and K. Johanson *et al.*, J Biol Chem, 270(16):9459-9471 (1995)).

25 The polynucleotides, polypeptides and antibodies to the polypeptide of the present invention may also be used to configure screening methods for detecting the effect of added compounds on the production of mRNA and polypeptide in cells. For example, an ELISA assay may be constructed for measuring secreted or cell associated levels of polypeptide using monoclonal and polyclonal antibodies by standard methods known in the art. This can be used to discover agents which may inhibit or enhance the production of polypeptide(also called antagonist or agonist, respectively) from suitably manipulated cells or tissues.

30 The polypeptide may be used to identify membrane bound or soluble receptors, if any, through standard receptor binding techniques known in the art. These include, but are not limited to, ligand binding and crosslinking assays in which the polypeptide is labeled with a radioactive isotope (for instance, ^{125}I), chemically modified (for instance, biotinylated), or fused to a peptide sequence suitable for detection or purification, and incubated with a source of the putative

receptor (cells, cell membranes, cell supernatants, tissue extracts, bodily fluids). Other methods include biophysical techniques such as surface plasmon resonance and spectroscopy. These screening methods may also be used to identify agonists and antagonists of the polypeptide which compete with the binding of the polypeptide to its receptors, if any. Standard methods for 5 conducting such assays are well understood in the art.

Examples of potential polypeptide antagonists include antibodies or, in some cases, oligonucleotides or proteins which are closely related to the ligands, substrates, receptors, enzymes, etc., as the case may be, of the polypeptide, e.g., a fragment of the ligands, substrates, receptors, enzymes, etc.; or small molecules which bind to the polypeptide of the present invention but do not 10 elicit a response, so that the activity of the polypeptide is prevented.

Thus, in another aspect, the present invention relates to a screening kit for identifying agonists, antagonists, ligands, receptors, substrates, enzymes, etc. for polypeptides of the present invention; or compounds which decrease or enhance the production of such polypeptides, which comprises:

15 (a) a polypeptide of the present invention;
(b) a recombinant cell expressing a polypeptide of the present invention;
(c) a cell membrane expressing a polypeptide of the present invention; or
(d) antibody to a polypeptide of the present invention;
which polypeptide is preferably that of SEQ ID NO:2 or SEQ ID NO:6.

20 It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component.

It will be readily appreciated by the skilled artisan that a polypeptide of the present invention may also be used in a method for the structure-based design of an agonist, antagonist or inhibitor of the polypeptide, by:

25 (a) determining in the first instance the three-dimensional structure of the polypeptide;
(b) deducing the three-dimensional structure for the likely reactive or binding site(s) of an agonist, antagonist or inhibitor;
(c) synthesizing candidate compounds that are predicted to bind to or react with the deduced binding or reactive site; and

30 (d) testing whether the candidate compounds are indeed agonists, antagonists or inhibitors.

It will be further appreciated that this will normally be an iterative process.

In a further aspect, the present invention provides methods of treating abnormal conditions such as, for instance, pain, chronic pain, neuropathic pain, postoperative pain, rheumatoid arthritic pain, neuralgia, neuropathies, algesia, nerve injury, ischaemia, neurodegeneration, stroke,

incontinence and inflammatory disorders, related to either an excess of, or an under-expression of, VANILREP2 polypeptide activity.

If the activity of the polypeptide is in excess, several approaches are available. One approach comprises administering to a subject in need thereof an inhibitor compound (antagonist) as hereinabove described, optionally in combination with a pharmaceutically acceptable carrier, in an amount effective to inhibit the function of the polypeptide, such as, for example, by blocking the binding of ligands, substrates, receptors, enzymes, etc., or by inhibiting a second signal, and thereby alleviating the abnormal condition. In another approach, soluble forms of the polypeptides still capable of binding the ligand, substrate, enzymes, receptors, etc. in competition with endogenous polypeptide may be administered. Typical examples of such competitors include fragments of the VANILREP2 polypeptide.

In still another approach, expression of the gene encoding endogenous VANILREP2 polypeptide can be inhibited using expression blocking techniques. Known such techniques involve the use of antisense sequences, either internally generated or externally administered (see, for example, O'Connor, J Neurochem (1991) 56:560 in Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Alternatively, oligonucleotides which form triple helices ("triplexes") with the gene can be supplied (see, for example, Lee *et al.*, Nucleic Acids Res (1979) 6:3073; Cooney *et al.*, Science (1988) 241:456; Dervan *et al.*, Science (1991) 251:1360). These oligomers can be administered *per se* or the relevant oligomers can be expressed *in vivo*. Synthetic antisense or triplex oligonucleotides may comprise modified bases or modified backbones. Examples of the latter include methylphosphonate, phosphorothioate or peptide nucleic acid backbones. Such backbones are incorporated in the antisense or triplex oligonucleotide in order to provide protection from degradation by nucleases and are well known in the art. Antisense and triplex molecules synthesised with these or other modified backbones also form part of the present invention.

In addition, expression of the human VANILREP2 polypeptide may be prevented by using ribozymes specific to the human VANILREP2 mRNA sequence. Ribozymes are catalytically active RNAs that can be natural or synthetic (see for example Usman, N, *et al.*, Curr. Opin. Struct. Biol (1996) 6(4), 527-33.) Synthetic ribozymes can be designed to specifically cleave human VANILREP2 mRNAs at selected positions thereby preventing translation of the human VANILREP2 mRNAs into functional polypeptide. Ribozymes may be synthesised with a natural ribose phosphate backbone and natural bases, as normally found in RNA molecules. Alternatively the ribozymes may be synthesised with non-natural backbones to provide protection from ribonuclease degradation, for example, 2'-O-methyl RNA, and may contain modified bases.

For treating abnormal conditions related to an under-expression of VANILREP2 and its activity, several approaches are also available. One approach comprises administering to a subject a therapeutically effective amount of a compound which activates a polypeptide of the present invention, i.e., an agonist as described above, in combination with a pharmaceutically acceptable carrier, to thereby alleviate the abnormal condition. Alternatively, gene therapy may be employed to effect the endogenous production of VANILREP2 by the relevant cells in the subject. For example, a polynucleotide of the invention may be engineered for expression in a replication defective retroviral vector, as discussed above. The retroviral expression construct may then be isolated and introduced into a packaging cell transduced with a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a subject for engineering cells *in vivo* and expression of the polypeptide *in vivo*. For an overview of gene therapy, see Chapter 20, Gene Therapy and other Molecular Genetic-based Therapeutic Approaches, (and references cited therein) in Human Molecular Genetics, T Strachan and A P Read, BIOS Scientific Publishers Ltd (1996). Another approach is to administer a therapeutic amount of a polypeptide of the present invention in combination with a suitable pharmaceutical carrier.

In a further aspect, the present invention provides for pharmaceutical compositions comprising a therapeutically effective amount of a polypeptide, such as the soluble form of a polypeptide of the present invention, agonist/antagonist peptide or small molecule compound, in combination with a pharmaceutically acceptable carrier or excipient. Such carriers include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention. Polypeptides and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

The composition will be adapted to the route of administration, for instance by a systemic or an oral route. Preferred forms of systemic administration include injection, typically by intravenous injection. Other injection routes, such as subcutaneous, intramuscular, or intraperitoneal, can be used. Alternative means for systemic administration include transmucosal and transdermal administration using penetrants such as bile salts or fusidic acids or other detergents. In addition, if a polypeptide or other compounds of the present invention can be formulated in an enteric or an encapsulated formulation, oral administration may also be possible. Administration of these compounds may also be topical and/or localized, in the form of salves, pastes, gels, and the like.

The dosage range required depends on the choice of peptide or other compounds of the present invention, the route of administration, the nature of the formulation, the nature of the subject's condition, and the judgment of the attending practitioner. Suitable dosages, however, are in the range of 0.1-100 µg/kg of subject. Wide variations in the needed dosage, however, are to be expected in view of the variety of compounds available and the differing efficiencies of various routes of administration. For example, oral administration would be expected to require higher dosages than administration by intravenous injection. Variations in these dosage levels can be adjusted using standard empirical routines for optimization, as is well understood in the art.

Polypeptides used in treatment can also be generated endogenously in the subject, in treatment modalities often referred to as "gene therapy" as described above. Thus, for example, cells from a subject may be engineered with a polynucleotide, such as a DNA or RNA, to encode a polypeptide *ex vivo*, and for example, by the use of a retroviral plasmid vector. The cells are then introduced into the subject.

Polynucleotide and polypeptide sequences form a valuable information resource with which to identify further sequences of similar homology. This is most easily facilitated by storing the sequence in a computer readable medium and then using the stored data to search a sequence database using well known searching tools, such as those in the GCG and Lasergene software packages. Accordingly, in a further aspect, the present invention provides for a computer readable medium having stored thereon a polynucleotide comprising the sequence of SEQ ID NO:1 or SEQ ID NO:5 and/or a polypeptide sequence encoded thereby.

The following definitions are provided to facilitate understanding of certain terms used frequently hereinbefore.

"Antibodies" as used herein includes polyclonal and monoclonal antibodies, chimeric, single chain, and humanized antibodies, as well as Fab fragments, including the products of an Fab or other immunoglobulin expression library.

"Isolated" means altered "by the hand of man" from the natural state. If an "isolated" composition or substance occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living animal is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein.

"Polynucleotide" generally refers to any polyribonucleotide or polydeoxyribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. "Polynucleotides" include, without limitation, single- and double-stranded DNA, DNA that is a mixture of single- and

double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, "polynucleotide" refers to triple-stranded regions comprising RNA or DNA or both 5 RNA and DNA. The term "polynucleotide" also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications may be made to DNA and RNA; thus, "polynucleotide" embraces 10 chemically, enzymatically or metabolically modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. "Polynucleotide" also embraces relatively short polynucleotides, often referred to as oligonucleotides.

"Polypeptide" refers to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres. 15 "Polypeptide" refers to both short chains, commonly referred to as peptides, oligopeptides or oligomers, and to longer chains, generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene-encoded amino acids. "Polypeptides" include amino acid sequences modified either by natural processes, such as post-translational processing, or by chemical modification techniques which are well known in the art. Such modifications are well 20 described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications may occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present to the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides 25 may be branched as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched and branched cyclic polypeptides may result from post-translation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, biotinylation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment 30 of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenylation, sulfation, transfer-RNA

mediated addition of amino acids to proteins such as arginylation, and ubiquitination (see, for instance, Proteins - Structure and Molecular Properties, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993; Wold, F., Post-translational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in Post-translational Covalent Modification of Proteins, B. C. Johnson, 5 Ed., Academic Press, New York, 1983; Seifter *et al.*, "Analysis for protein modifications and nonprotein cofactors", *Meth Enzymol* (1990) 182:626-646 and Rattan *et al.*, "Protein Synthesis: Post-translational Modifications and Aging", *Ann NY Acad Sci* (1992) 663:48-62).

"Variant" refers to a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide, but retains essential properties. A typical variant of a 10 polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid 15 sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide 20 or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis.

"Identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences.

25 In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" and "similarity" can be readily calculated by known methods, including but not limited to those described in (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, 30 Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data*, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., *SIAM J. Applied Math.*, 48: 1073 (1988). Preferred methods to determine identity

are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, the GCG program package (Devereux, J., et al., *Nucleic Acids Research* 12(1): 387 (1984)), BLASTP, BLASTN, and FASTA (Altschul, S.F. et al., *J. Molec. Biol.* 215: 403-410 (1990). The BLAST X program is publicly available from NCBI and other sources (*BLAST Manual*, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., et al., *J. Mol. Biol.* 215: 403-410 (1990). The well known Smith Waterman algorithm may also be used to determine identity.

10 Preferred parameters for polypeptide sequence comparison include the following:

1) Algorithm: Needleman and Wunsch, *J. Mol Biol.* 48: 443-453 (1970)

Comparison matrix: BLOSSUM62 from Hentikoff and Hentikoff, *Proc. Natl. Acad. Sci. USA.* 89:10915-10919 (1992)

Gap Penalty: 12

15 Gap Length Penalty: 4

A program useful with these parameters is publicly available as the "gap" program from Genetics Computer Group, Madison WI. The aforementioned parameters are the default parameters for peptide comparisons (along with no penalty for end gaps).

Preferred parameters for polynucleotide comparison include the following:

20 1) Algorithm: Needleman and Wunsch, *J. Mol Biol.* 48: 443-453 (1970)

Comparison matrix: matches = +10, mismatch = 0

Gap Penalty: 50

Gap Length Penalty: 3

Available as: The "gap" program from Genetics Computer Group, Madison WI. These are the 25 default parameters for nucleic acid comparisons.

By way of example, a polynucleotide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:1, that is be 100% identical, or it may include up to a certain integer number of nucleotide alterations as compared to the reference sequence. Such alterations are selected from the group consisting of at least one nucleotide deletion, substitution, 30 including transition and transversion, or insertion, and wherein said alterations may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence. The number of nucleotide alterations is determined by multiplying the total number of nucleotides in SEQ ID NO:1 by the

numerical percent of the respective percent identity(divided by 100) and subtracting that product from said total number of nucleotides in SEQ ID NO:1, or:

$$n_n \leq x_n - (x_n \cdot y),$$

wherein n_n is the number of nucleotide alterations, x_n is the total number of nucleotides in SEQ

5 ID NO:1, and y is, for instance, 0.70 for 70%, 0.80 for 80%, 0.85 for 85%, 0.90 for 90%, 0.95 for 95%,etc., and wherein any non-integer product of x_n and y is rounded down to the nearest integer prior to subtracting it from x_n . Alterations of a polynucleotide sequence encoding the polypeptide of SEQ ID NO:2 may create nonsense, missense or frameshift mutations in this coding sequence and thereby alter the polypeptide encoded by the polynucleotide following such 10 alterations.

Similarly, a polypeptide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:2, that is be 100% identical, or it may include up to a certain integer number of amino acid alterations as compared to the reference sequence such that the %

identity is less than 100%. Such alterations are selected from the group consisting of at least one 15 amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and wherein said alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence. The number of amino acid alterations for a given % 20 identity is determined by multiplying the total number of amino acids in SEQ ID NO:2 by the numerical percent of the respective percent identity(divided by 100) and then subtracting that product from said total number of amino acids in SEQ ID NO:2, or:

$$n_a \leq x_a - (x_a \cdot y),$$

wherein n_a is the number of amino acid alterations, x_a is the total number of amino acids in SEQ

25 ID NO:2, and y is, for instance 0.70 for 70%, 0.80 for 80%, 0.85 for 85% etc., and wherein any non-integer product of x_a and y is rounded down to the nearest integer prior to subtracting it from x_a .

"Homolog" is a generic term used in the art to indicate a polynucleotide or polypeptide sequence possessing a high degree of sequence relatedness to a subject sequence. Such

30 relatedness may be quantified by determining the degree of identity and/or similarity between the sequences being compared as hereinbefore described. Falling within this generic term are the terms "ortholog", meaning a polynucleotide or polypeptide that is the functional equivalent of a

polynucleotide or polypeptide in another species, and "paralog" meaning a functionally similar sequence when considered within the same species.

"Fusion protein" refers to a protein encoded by two, often unrelated, fused genes or fragments thereof. In one example, EP-A-0 464 discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, employing an immunoglobulin Fc region as a part of a fusion protein is advantageous for use in therapy and diagnosis resulting in, for example, improved pharmacokinetic properties [see, e.g., EP-A 0232 262]. On the other hand, for some uses it would be desirable to be able to delete the Fc part after the fusion protein has been expressed, detected and purified.

Example

Analysis of VANILREP2 mRNA expression was assessed during monocyte to macrophage maturation and in cells of the mononuclear phagocyte cell lineage in response to an inflammatory

mediator such as might be involved in hyperalgesia.

Monocytes were prepared by counter-current centrifugal elutriation using a method adapted from Stevenson et al (1983). Mononuclear white blood cells were purified using density gradient

centrifugation through Histopaque. The collected cells were resuspended in PBS w/o calcium or magnesium and supplemented with 2% autologous serum and loaded into the elutriator chamber.

The counterflow rate was increased incrementally and monocytes collected when their purity reached 95%, typically at a flow rate of 18ml/min. The monocytes were then cultured in RPMI supplemented with 2% human serum. C13 cells were cultured in DMEM supplemented with glutamine and 10% fetal calf serum and treated with either vehicle or IL-1alpha at 1ng/ml for 24hrs.

At the appropriate times the cells were harvested using Trizol and RNA isolated according to the manufacturers instructions. RNA was reverse transcribed using SuperscriptII and the resulting cDNA used in a Taqman assay. Sample cDNA and appropriate no RT controls were added to a 25 microL final reaction volume containing Taqman primers/probe and were run on a Perkin Elmer 7700 Taqman machine for 40 cycles. Three different housekeeping genes were measured in the

samples, acidic ribosomal phosphoprotein PO (36B4), hypoxanthine-guanine phosphoribosyl transferase (HPRT) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH). Each transcript of interest was normalised to all three housekeeping genes.

Stevenson, H.C., et al. J. Immun. Methods 62 (1983) 353-363.

Expression of VANILREP2 was found to be upregulated in monocytes during the course of their maturation into macrophages after four days. VANILREP2 was also upregulated in C13 cells following exposure to 1ng/ml IL-1alpha for 24hours.

5 All publications, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference as if each individual publication were specifically and individually indicated to be incorporated by reference herein as though fully set forth.

SEQUENCE INFORMATION

SEQ ID NO:1

TAGGATGACCTCACCTCCAGCTCTCAGTTTCAGGTTGGAGACATTAGATGGAGGCCA
 AGAAGATGGCTCTGAGGCGGACAGAGGAAAGCTGGATTTGGAGCGGGCTGCCTCCAT
 5 GGAGTCACAGTCCAGGGCGAGGACCGAAATTGCCCTCAGATAAGAGTCAACCTCAA
 CTACCGAAAGGAAACAGGTGCCAGTCAGCCGATCCAAACCGATTGACCGAGATCGGCT
 CTTCAATGCGGTCTCCGGGTGTCCCCGAGGATCTGGCTGGACTTCCAGAGTACCTGAG
 CAAGACCAGCAAGTACCTCACCGACTCGGAATACACAGAGGGCTCCACAGGTAAGACGTG
 CCTGATGAAGGCTGTGCTAACCTAACGGACGGGTCAATGCCCTGCATTCTGCCACTGCT
 10 GCAGATCGACCGGGACTCTGGCAATCCTCAGCCCTGGTAAATGCCAGTGCACAGATGA
 CTATTACCGAGGCCACAGCGCTCTGCACATGCCATTGAGAAGAGGAGTCTGCAGTGTGT
 GAAGCTCCTGGTGGAGAATGGGCCAATGTGCATGCCCGGCTGCCGCGCTTCTTCCA
 GAAGGCCAAGGGACTGCTTTATTCGGTGAGCTACCCCTCTTTGCCGCTTGCAC
 15 CAAGCAGTGGATGTGTAAGCTACCTCCTGGAGAACCCACACCAGCCGCCAGCCTGCA
 GGCCACTGACTCCCAGGGCAACACAGTCCTGCATGCCCTAGTGATGATCTCGGACAACTC
 AGCTGAGAACATTGCACTGGTACCAGCATGTATGATGGCTCCTCAAAGCTGGGCCCG
 CCTCTGCCCTACCGTGCAGCTTGAGGACATCCGAACCTGCAGGATCTCACGCCCTGAA
 GCTGGCCGCCAAGGAGGGCAAGATCGAGATTTCAAGGCACATCCTGCAGCGGGAGTTTC
 AGGACTGAGCCACCTTCCGAAAGTTCACCGAGTGGTGCTATGGGCCCTGCCGGGTGTC
 20 GCTGTATGACCTGGCTCTGTGGACAGCTGTGAGGAGAACTCAGTGCTGGAGATCATTGC
 CTTTCATTGCAAGAGCCCGCACCGACACCGAATGGTCGTTGGAGGCCCTGAACAAACT
 GCTGCAGGCAGAAATGGGATCTGCTCATCCCCAAGTTCTTAAACTTCTGTGTAATCT
 GATCTACATGTTCATCTCACCGCTGTGCCTACCATCAGCCTACCCCTGAAGAAGCAGGC
 CGCCCTCACCTGAAAGCGGAGGTTGAAACTCCATGCTGCTGACGGGCCACATCCTTAT
 25 CCTGCTAGGGGGATCTACCTCCTCGTGGCCAGCTGTGGTACTTCTGGCGGCCACGT
 GTTCATCTGGATCTGTTCATAGACAGCTACTTTGAAATCCTCTTCCCTGTTCCAGGCCCT
 GCTCACAGTGGTGTCCCAGGTGCTGTGTTCTGCCATCGAGTGGTACCTGCCCTGCT
 TGTGTCTGCGCTGGTCTGGCTGGCTGAACCTGCTTACTATACACGTGGCTTCAGCA
 CACAGGCATCTACAGTGTATGATCCAGAAGGTATCCTGCCGACCTGCTGCGCTTCCCT
 30 TCTGATCTACTTAGTCCTTTCGGCTTCGCTGTAGCCCTGGTGAGCCTGAGCCAGGA
 GGCTTGGCGCCCCGAAGCTCCTACAGGCCCAATGCCACAGAGTCAGTGCAAGCCATGGA
 GGGACAGGGAGGACGAGGGCAACGGGGCCAGTACAGGGTATCTGGAAGCCTCTGG
 GCTCTCAAATTACCATCGGCATGGCGAGCTGGCCTTCCAGGAGCAGCTGCACTTCCG
 CGGCATGGTGTGCTGCTGCTGGCCTACGTGCTGCTCACCTACATCCTGCTGCTCAA
 35 CATGCTCATGCCCTCATGAGCGAGACCGTCAACAGTGTGCCACTGACAGCTGGAGCAT
 CTGGAAGCTGCAGAAAGCCATCTGTCTGGAGATGGAGAATGGCTATTGGTGGTGCAG
 GAAGAAGCAGCGGGCAGGTGTGATGCTGACCGTTGCCACTAAGCCAGATGGCAGCCCGGA

TGAGCGCTGGTCTTCAGGGTGGAGGAGGTGAACCTGGGCTTCATGGGAGCAGACGCTGCC
 TACGCTGTGAGGACCCGTCAAGGGCAGGTGTCCTCGAACACTCTCGAGAACCCGTCC
 GGCTTCCCCTCCCAAGGAGGATGAGGATGGTGCCTCTGAGGAAAACATATGTGCCGTCCA
 GCTCCTCCAGTCCAACGTGATGGCCCAGATGCAGCAGGAGGCCAGAGGACAGAGCAGAGGA
 5 TCTTTCCAACC

SEQ ID NO:2

MTSPSSSPVFRLETLDGGQEDGSEADRGKLDGSGLPPMESQFQGEDRKFAPQIRVNLY
 RKGTGASQPDPNRFDRDRLFNAVRGVPEDLAGLPEYLSKTSKYLTDSEYTEGSTGKTCL
 10 MKAVLNLDGVNACILPLLQIDRDSGNPQPLVNAQCTDDYYRGHSALHIAIEKRSQCVK
 LLVENGANVHARACGRFFQKGQGTCFYFGEPLSLAACTKQWDVVSYLLENPHQPASLQA
 TDSQGNTVLHALVMISDNAENIALVTSMYDGLLQAGARLCPTVQLEDIRNLQDLTPLKL
 AAKEGKIEIFRHILQREFSGLSHLSRKFTEWCYGPVRVSLYDLASVDSCEENSVLEIIAF
 HCKSPHRHRMVMVLEPLNKLQQAKWDLLIPKFFLNFLCNLIYMFIFTAVAYHQPTLKKQAA
 15 PHLKAEVGNSMLLTGHILILLGGIYLLVQLWYFWRRHVFIIWISFIDSYFEILFLFQALL
 TVVSQVLCLFLAIEWYLPPLLVSALVLGWLNLYYTRGFQHTGIYSVMIQKVLRDLLRFLL
 IYLVFLFGFAVALVSLSQEAWRPEAPTGPNATESVQPMEGQEDEGNGAQYRGILEASLEL
 FKFTIGMGEAFQEQQLHFRGMVLLLLAYVLLTYILLNMLIALMSETVNSVATDSWSIW
 KLQKAISVLEMENGYWWCRKKQRAGVMLTVGTPDGPDERWCFRVEEVNWASWEQTLPT
 20 LCEDPSGAGVPRTLENPVLASPPKEDEDGASEENYVPVQQLQSN

SEQ ID NO:3

TGTTTCTGGCCATCGAGTGGTACCTGCCCTGCTTGTCTGCGCTGGTCTGGGCTGG
 25 CTGAACCTGCTTACTATACACGTGGCTCCAGCACACAGGCATCTACAGTGTCTGATC
 CAGAACCCCTGGTGAGCCTGAGCCAGGANNTGGCGCCCGAAGCTCTACAGGCCCAA
 TGCCACAGAGTCAGTGCAGCCATGGAGGGACAGGAGGACGAGGGCAACGGGGCCAGTA
 CAGGGGTATCCTGGAAAGCCTCTGGAGCTCTCAAATTCAACCATCGGCATGGCGAGCT
 GGCCTTCAGGAGCAGCTGCACTCCGGCATGGTCTGCTGCTGCTGGCCTACGTC
 30 GCTGCTCACCTACATCCTGCTCAACATGCTCATGCCCTCATGAGCGAGACCGTCAA
 CAGTGTGCCACTGACAGCTGGAGCATCTGGAAGCTGCAGAAAGCNATCTNTGCTGG
 GATGGAGAATGGCTATTGGTGGTGCAGGAAGAAGCAGCGGGCAGGTGTGATGCTGACCGT
 TGGCACTAAGCCAGATGGCAGGCCGATGAGCGCTGGTCTCAGGGTGGAGGAGGTGAA
 CTGGGCTTCATGGGAGCAGACGCTGCCACGCTGCTGAGGACCCGTCAAGGGCAGGTGT
 35 CCCTCGAACTCTCGAGAACCCCTGTCCTGGCTTCCCCTCCCAAGGAGGATGAGGATGGTGC
 CTCTGAGGAAAACATATGTGCCGTCCAGCTCTCCAGTCCAACGTGATGGCCCAGATGCAG
 CAGGAGGCCAGAGGACAGAGCAGAGGATCTTCCAACCACATCTGCTGGCTCTGGGTCC

CAGTGAATTCTGGTGGCAAATATATATTTCACTAACTCAAAAC

SEQ ID NO:4

5 Smith-Waterman score: 579; 43.8% identity in 242 aa overlap

Query: YPLLVSALVLGWLNLYYTRGFQHTGIYSVMIQKPW-----XAXA

Sbjct: YVASMVFSLAMGWTNMLYYTRGFQQMGIYAVMIEKMLRDLCRFMFVYLVFLFGFSTAVV
540 550 560 570 580 590

Query: RKLAP-RSSYRPQCHRVSAAHGGTG\DEGNGAQYRGILEASLELFKFTIGMGEAFQEQL

Sbjct: TLIEDGKNNSLPMESTPHKCRGSAC-KPGNS--YNSLYSTCLELFKFTIGMDLEFTENY
600 610 620 630

Query: HFRGMVLLLLAYVLLTYILLLNMLIALMSETVNSVATDSWSIWKLQKAIYVLEMENGYW

Sbjct: DFKAVFIILLAYVILTYILLNMLIALMGETVNKIAQESKNIWKLQRAITILDTEKSFL
668

Query: --WCR-KKQRAGVMLTVGTPDGSPDERWCFRVEEVNWASWEQTLPTLCEDPSGA-GVPR

Sbjct: --KCMRKA FRSGKLLQVGFTP DGKDDYRWC RVDEVNWTWNTNVGI INEDPGNCEGVKR

Query: TI

•

Subject: TL

30 In the foregoing, SEQ ID NO:4 refers to any one or more of the sequences designated by the symbol Query. Symbol Sbjct refers to the reference sequence, rat vanilloid receptor, VR1 (AF029310)

SEQ ID NO:5 (Polymorphic variant PVP-1 of VANU BPF2)

35 TAGGATGACCTCACCTCCAGCTCTCCAGTTTCAGGTTGGAGACATTAGATGGAGGCCA
AGAAGATGGCTCTGAGGCGGACAGAGGAAAGCTGGATTTGGGAGCGGGCTGCCTCCCAT
GGAGTCACAGTTCCAGGGCGAGGACCGGAAATCGCCCCCTCAGATAAGAGTCACCTCA

CTACCGAAAGGGAACAGGTGCCAGTCAGCCGGATCCAAACCGATTGACCGAGATCGGCT
CTTCAATGCGGTCTCCGGGGTGTCCCCGAGGATCTGGCTGGACTTCCAGAGTACCTGAG
CAAGACCAAGTACCTCACCGACTCGGAATACACAGAGGGCTCCACAGGTAAGACGTG
CCTGATGAAGGCTGTGCTGAACCTTAAGGACGGAGTCAATGCCCTGCATTCTGCCACTGCT
5 GCAGATCGACAGGGACTCTGGCAATCCTCAGCCCCCTGGTAAATGCCCAAGTGCACAGATGA
CTATTACCGAGGCCACAGCGCTCTGCACATGCCATTGAGAAGAGGGAGTCTGCAGTGTGT
GAAGCTCCTGGTGGAGAATGGGGCAATGTGCATGCCCGGCTGCCGCTTCTTCCA
GAAGGGCCAAGGGACTTGCTTTATTCGGTGAGCTACCCCTCTTTGGCCGCTTGCAC
CAAGCAGTGGGATGTGTTAAGCTACCTCTGGAGAACCCACACCAGCCCCCAGCCTGCA
10 GGCCACTGACTCCCAGGGCAACACAGTCCTGCATGCCCTAGTGATGATCTCGGACAACTC
AGCTGAGAACATTGCACTGGTGACCAGCATGTATGATGGCTCCTCCAAGCTGGGGCCCG
CCTCTGCCCTACCGTGCAGCTTGAGGACATCGCAACCTGCAGGATCTCACGCCCTGAA
GCTGGCCGCCAAGGAGGGCAAGATCGAGATTTCAAGGCACATCCTGCAGCGGGAGTTTC
AGGACTGAGCCACCTTCCGAAAGTTCACCGAGTGGTGCTATGGGCTGTCCGGGTGTC
15 GCTGTATGACCTGGCTTCTGTGGACAGCTGTGAGGAGAACTCAGTGCTGGAGATCATTG
CTTTCATTGCAAGAGCCCGCACCGACACCGAATGGTCGTTGGAGCCCTGAACAAACT
GCTGCAGGCAGGAAATGGGATCTGCTCATCCCCAAGTTCTTAAACTCCTGTGTAATCT
GATCTACATGTTCATCTCACCGCTGTTGCCTACCATCAGCCTACCCCTGAAGAAGGCCGC
CCCTCACCTGAAAGCGGAGGTTGGAAACTCCATGCTGCTGACGGGCCACATCCTTATCCT
20 GCTAGGGGGGATCTACCTCCTCGTGGCCAGCTGTGGTACTTCTGGCGGCCACGTGTT
CATCTGGATCTCGTCATAGACAGCTACTTGAAATCCTCTTCCCTGTTCCAGGCCCTGCT
CACAGTGGTGTCCCAGGTGCTGTGTTCCCTGGCCATCGAGTGGTACCTGCCCTGCTTGT
GTCTGCGCTGGTGCTGGCTGGCTGAACCTGCTTACTATACACGTGGCTCCAGCACAC
AGGCATCTACAGTGTGATCCAGAAGGTGATCCTGCAGGGACCTGCTGCGCTTCTTCT
25 GATCTACTTAGTCTTCCTTTCGGCTTCGCTGTAGCCCTGGTGAGCCTGAGCCAGGAGGC
TTGGCGCCCGAAGCTCCTACAGGCCCAATGCCACAGAGTCAGTGCAAGCCATGGAGGG
ACAGGAGGACGAGGGCAACGGGGCCAGTACAGGGGTATCCTGGAAGCCTCTGGAGCT
CTTCAAATTCAACCATCGGATGGCGAGCTGGCCTCCAGGAGCAGCTGCACTCCGCGG
CATGGTGCTGCTGCTGCTGGCCTACGTGCTGCTCACCTACATCCTGCTGCTCAACAT
30 GCTCATGCCCTCATGAGCGAGACCGTCAACAGTGTGCCACTGACAGCTGGAGCATTG
GAAGCTGCAGAAAGCCATCTCTGCTGGAGATGGAGAATGGCTATTGGTGGTGAGGAA
GAAGCAGCGGGCAGGTGTGATGCTGACCGTTGGCACTAAGCCAGATGGCAGCCCCGATGA
GCGCTGGTGCTTCAGGGTGGAGGGAGGTGAACCTGGCTTACATGGGAGCAGCCTGCTAC
GCTGTGTGAGGACCCGTCAAGGGCAGGTGCTCCCTGAACTCTCGAGAACCCTGTCTGGC
35 TTCCCCCTCCAAGGAGGATGAGGATGGTGCTCTGAGGAAAACATGTGCCCCGTCCAGCT
CCTCCAGTCCAACGTGATGGCCAGATGCAGCAGGAGGCCAGAGGACAGAGCAGAGGATCT
TTCCAACC

SEQ ID NO:6 (polymorphic variant PVP-1 of VANILREP2)

MTSPSSSPVFRLETLDGGQEDGSEADRGKLDGSGLPPMESQFQGEDRKFAPQIRVNLNY
RKGTGASQPDPNRFDRDRLFNAVRGVPEDLAGLPEYLSKTSKYLTDSEYTEGSTGKTCL
5 MKAVLNLKDGVNACILPLLQIDRDSGNPQPLVNAQCTDDYYRGHSALHIAIEKRSLQCVK
LLVENGANVHARACGRFFQKGQGTCFYFGELPLSLAACTKQWDVVSYLLENPHQPASLQA
TDSQGNTVLHALVMISDNSAENIALVTSMYDGLLQAGARILCPTVQLEDIRNLQDLTPLKL
AAKEGKIEIFRHILQREFSGLSHLSRKFTEWCYGPVRVSLYDLASVDSCEENSVLEIIAF
10 HCKSPHRHRMVVLEPLNKLLQAKWDLLIPKFFLNFLCNLIYMFIFTAVAYHQPTLKKAAP
YLVFLFGFAVALVSLSQEAWRPEAPTGPNATESVQPMEGQEDEGNGAQYRGILEASLELF
KFTIGMGECAFQEQLHFRGMVLLLLAYVLLTYILLNMLIALMSETVNSVATDSWSIWK
15 LQKAISVLEMENGYWWCRKKQRAGVMLTVGTPDGSPDERWCFRVEEVNWASWEQTLPTL
CEDPSGAGVPRTLENPVLASPPKEDEDGASEENYVPVQLLQSN

SEQUENCE LISTING

5 <110> SmithKline Beecham plc

10 <120> Novel Compounds

15 <130> GP30040

20 <160> 6

25 <170> FastSEQ for Windows Version 3.0

30 <210> 1

35 <211> 2351

40 <212> DNA

45 <213> Homo sapiens

50 <400> 1

55	taggatgacc tcaccctcca gctctccagt tttcaggttg gagacattag atggaggcca	60
	agaagatggc tctgaggcgg acagagggaa gctggattt gggagcgggc tgcctccat	120
	ggagtcacag ttccagggcg aggaccgaa attcgcccc cagataagag tcaacctcaa	180
	ctaccgaaaag ggaacaggtg ccagtcagcc ggatccaaac cgatttacc gagatcggt	240
60	cttcaatgcg gtctccggg gtgtcccgaa ggatctggct ggacttccag agtacctgag	300
	caagaccagc aagtacctca ccgactcggg atacacagag ggctccacag gtaagacgtg	360
	cctgtatgaag gctgtgctga accttaaggg cggggtaat gcctgcattc tgccactgt	420
	gcagatcgac cgggactctg gcaatcctca gcccctggta aatgcccagt gcacagatga	480
65	ctattaccga ggcacacagcg ctctgcacat cgcatttggag aagaggagtc tgcagtgtgt	540
	gaaggtctcg gtggagaatg gggccaaatgt gcatgcccgg gcctgcggcc gcttcttcca	600
	gaaggccaa gggacttgct tttatattcg tgagctaccc ctctctttgg cgccttgcac	660
	caagcagtgg gatgtggtaa gctacctcct ggagaaccca caccagcccg ccagctgca	720
70	ggccactgac tcccagggca acacagtccct gcatgccta gtgatgtatct cggacaactc	780
	agctgagaac attgcactgg tgaccagcat gtatgtgg ctccctccaag ctggggcccg	840
	cctctgcctt accgtgcagc ttgaggacat ccgcaacctg caggatctca cgcctctgaa	900
	gctggccccc aaggaggggca agatcgagat tttcaggcac atcctgcagc gggagtttc	960
	aggactgagc cacctttccc gaaagttcac cgagtggtgc tatggccctg tccgggtgtc	1020
	gctgtatgac ctggcttcgt tggacagctg tgaggagaac tcaagtgttgg agatcattgc	1080
75	ctttcatgc aagagccccgc accgacacccg aatggtcgtt ttggagcccc tgaacaaact	1140
	gctgcaggcg aaatgggatc tgctcatccc caagtttttc ttaaacttcc tggtaatct	1200
	gatctacatg ttcatcttca ccgctgttgc ctaccatcag cctaccctga agaaggcaggc	1260

20 2331

<210> 2

<211> 764

<212> PRT

<213> *Homo sapiens*

25

<400> ?

Met Thr Ser Pro Ser Ser Ser Pro Val Phe Arg Ile Glu

1 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 Ser Thr Val Phe Arg Leu Glu Thr Leu

1 5 10 15
Gly Gly Gln Glu Asp Gly Ser Glu Ala Asp Arg Gly Lys Leu Asp Phe
30 20 25 30

15 30

Gly Ser Gly Leu Pro Pro Met Glu Ser Gln Phe Gln Gly Glu Asp Arg

35	40	45
Lys Phe Ala Pro Gln Ile Arg Val Asn Leu Asn Tyr Arg Lys Gly Thr		
50	55	60

Gly-Ala-Ser-Gly-Der- β -D-Glucosidase

Ash Ala Val Ser Arg Gly Val Pro Glu Asp Leu Ala Gly Leu Pro Glu

85 90 95
Tyr Leu Ser Lys Thr Ser Lys Tyr Leu Thr Asp Ser Glu Tyr Thr Glu

101 105 110
Gly Ser Thr Gly Lys Thr Cys Leu Met Lys Ala Val Leu Asp Leu Lys

	115	120	125
	Asp Gly Val Asn Ala Cys Ile Leu Pro Leu Leu Gln Ile Asp Arg Asp		
	130	135	140
	Ser Gly Asn Pro Gln Pro Leu Val Asn Ala Gln Cys Thr Asp Asp Tyr		
5	145	150	155
	Tyr Arg Gly His Ser Ala Leu His Ile Ala Ile Glu Lys Arg Ser Leu		
	165	170	175
	Gln Cys Val Lys Leu Leu Val Glu Asn Gly Ala Asn Val His Ala Arg		
	180	185	190
10	Ala Cys Gly Arg Phe Phe Gln Lys Gly Gln Gly Thr Cys Phe Tyr Phe		
	195	200	205
	Gly Glu Leu Pro Leu Ser Leu Ala Ala Cys Thr Lys Gln Trp Asp Val		
	210	215	220
	Val Ser Tyr Leu Leu Glu Asn Pro His Gln Pro Ala Ser Leu Gln Ala		
15	225	230	235
	Thr Asp Ser Gln Gly Asn Thr Val Leu His Ala Leu Val Met Ile Ser		
	245	250	255
	Asp Asn Ser Ala Glu Asn Ile Ala Leu Val Thr Ser Met Tyr Asp Gly		
	260	265	270
20	Leu Leu Gln Ala Gly Ala Arg Leu Cys Pro Thr Val Gln Leu Glu Asp		
	275	280	285
	Ile Arg Asn Leu Gln Asp Leu Thr Pro Leu Lys Leu Ala Ala Lys Glu		
	290	295	300
	Gly Lys Ile Glu Ile Phe Arg His Ile Leu Gln Arg Glu Phe Ser Gly		
25	305	310	315
	Leu Ser His Leu Ser Arg Lys Phe Thr Glu Trp Cys Tyr Gly Pro Val		
	325	330	335
	Arg Val Ser Leu Tyr Asp Leu Ala Ser Val Asp Ser Cys Glu Glu Asn		
	340	345	350
30	Ser Val Leu Glu Ile Ile Ala Phe His Cys Lys Ser Pro His Arg His		
	355	360	365
	Arg Met Val Val Leu Glu Pro Leu Asn Lys Leu Leu Gln Ala Lys Trp		
	370	375	380
	Asp Leu Leu Ile Pro Lys Phe Phe Leu Asn Phe Leu Cys Asn Leu Ile		
35	385	390	395
	Tyr Met Phe Ile Phe Thr Ala Val Ala Tyr His Gln Pro Thr Leu Lys		
	405	410	415
	Lys Gln Ala Ala Pro His Leu Lys Ala Glu Val Gly Asn Ser Met Leu		
	420	425	430
40	Leu Thr Gly His Ile Leu Ile Leu Gly Gly Ile Tyr Leu Leu Val		
	435	440	445

Gly Gln Leu Trp Tyr Phe Trp Arg Arg His Val Phe Ile Trp Ile Ser
 450 455 460
 Phe Ile Asp Ser Tyr Phe Glu Ile Leu Phe Leu Phe Gln Ala Leu Leu
 465 470 475 480
 5 Thr Val Val Ser Gln Val Leu Cys Phe Leu Ala Ile Glu Trp Tyr Leu
 485 490 495
 Pro Leu Leu Val Ser Ala Leu Val Leu Gly Trp Leu Asn Leu Leu Tyr
 500 505 510
 Tyr Thr Arg Gly Phe Gln His Thr Gly Ile Tyr Ser Val Met Ile Gln
 10 515 520 525
 Lys Val Ile Leu Arg Asp Leu Leu Arg Phe Leu Leu Ile Tyr Leu Val
 530 535 540
 Phe Leu Phe Gly Phe Ala Val Ala Leu Val Ser Leu Ser Gln Glu Ala
 545 550 555 560
 15 Trp Arg Pro Glu Ala Pro Thr Gly Pro Asn Ala Thr Glu Ser Val Gln
 565 570 575
 Pro Met Glu Gly Gln Glu Asp Glu Gly Asn Gly Ala Gln Tyr Arg Gly
 580 585 590
 Ile Leu Glu Ala Ser Leu Glu Leu Phe Lys Phe Thr Ile Gly Met Gly
 20 595 600 605
 Glu Leu Ala Phe Gln Glu Gln Leu His Phe Arg Gly Met Val Leu Leu
 610 615 620
 Leu Leu Leu Ala Tyr Val Leu Leu Thr Tyr Ile Leu Leu Leu Asn Met
 625 630 635 640
 25 Leu Ile Ala Leu Met Ser Glu Thr Val Asn Ser Val Ala Thr Asp Ser
 645 650 655
 Trp Ser Ile Trp Lys Leu Gln Lys Ala Ile Ser Val Leu Glu Met Glu
 660 665 670
 Asn Gly Tyr Trp Trp Cys Arg Lys Lys Gln Arg Ala Gly Val Met Leu
 675 680 685
 30 Thr Val Gly Thr Lys Pro Asp Gly Ser Pro Asp Glu Arg Trp Cys Phe
 690 695 700
 Arg Val Glu Glu Val Asn Trp Ala Ser Trp Glu Gln Thr Leu Pro Thr
 705 710 715 720
 35 Leu Cys Glu Asp Pro Ser Gly Ala Gly Val Pro Arg Thr Leu Glu Asn
 725 730 735
 Pro Val Leu Ala Ser Pro Pro Lys Glu Asp Glu Asp Gly Ala Ser Glu
 740 745 750
 Glu Asn Tyr Val Pro Val Gln Leu Leu Gln Ser Asn
 40 755 760

<210> 3
 <211> 885
 <212> DNA
 <213> Homo sapiens

5

<400> 3

tgtttcctgg	ccatcgagtg	gtacctgccc	ctgcttgtgt	ctgcgctgg	gctgggctgg	60	
ctgaacctgc	tttactatac	acgtggcttc	cagcacacag	gcatctacag	tgtcatgatc	120	
cagaagccct	ggttagccctg	agccaggann	ttggcgcccc	gaagctccct	caggccccaa	180	
10	tgccacacag	tcagtgcagc	ccatggaggg	acaggaggac	gagggcaacg	240	
caggggtatc	ctggaagcct	ccttggagct	cttcaaattc	accatggca	tgggcgagct	300	
ggccttccag	gagcagctgc	acttccgcgg	catggtgctg	ctgctgtgc	tggcttacgt	360	
gctgctcacc	tacatcctgc	tgctcaacat	gctcatcgcc	ctcatgagcg	agaccgtcaa	420	
15	cagtgtcgcc	actgacagct	ggagcatctg	gaagctgcag	aaagcnatct	ntgtcctgga	480
gatggagaat	ggctatttgtt	ggtgcaggaa	gaagcagcgg	gcaggtgtga	tgctgaccgt	540	
tggcaactaag	ccagatggca	gccccgatga	gcgcgtggc	ttcagggtgg	aggaggtgaa	600	
ctgggcttca	tgggagcaga	cgctgcctac	gctgtgtgag	gaccgcgtcag	gggcaggtgt	660	
ccctcgaact	ctcgagaacc	ctgttctggc	ttccctcccc	aaggaggatg	aggatgggtgc	720	
20	ctctgaggaa	aactatgtgc	ccgtccagct	cctccagtc	aactgatggc	ccagatgcag	780
caggaggcca	gaggacagag	cagaggatct	ttccaaccac	atctgctggc	tctgggtcc	840	
cagtgaattc	tggtggcaaa	tatataattt	caactca	aaaac		885	

<210> 4

<211> 217

25

<212> PRT

<213> Homo sapiens

<400> 4

Tyr	Leu	Pro	Leu	Leu	Val	Ser	Ala	Leu	Val	Leu	Gly	Trp	Leu	Asn	Leu	
30	1							10					15			
Leu	Tyr	Tyr	Thr	Arg	Gly	Phe	Gln	His	Thr	Gly	Ile	Tyr	Ser	Val	Met	
									20				25		30	
Ile	Gln	Lys	Pro	Trp	Xaa	Ala	Xaa	Ala	Arg	Lys	Leu	Ala	Pro	Arg	Ser	
								35				40		45		
35	Ser	Tyr	Arg	Pro	Gln	Cys	His	Arg	Val	Ser	Ala	Ala	His	Gly	Thr	
								50				55		60		
Gly	Asp	Glu	Gly	Asn	Gly	Ala	Gln	Tyr	Arg	Gly	Ile	Leu	Glu	Ala	Ser	
							65				70		75		80	
40	Leu	Glu	Leu	Phe	Lys	Phe	Thr	Ile	Gly	Met	Gly	Glu	Leu	Ala	Phe	Gln
							85			90			95			
Glu	Gln	Leu	His	Phe	Arg	Gly	Met	Val	Leu	Leu	Leu	Leu	Ala	Tyr		

	100	105	110
	Val Leu Leu Thr Tyr Ile Leu Leu Leu Asn Met Leu Ile Ala Leu Met		
	115	120	125
	Ser Glu Thr Val Asn Ser Val Ala Thr Asp Ser Trp Ser Ile Trp Lys		
5	130	135	140
	Leu Gln Lys Ala Ile Tyr Val Leu Glu Met Glu Asn Gly Tyr Trp Trp		
	145	150	155
	Cys Arg Lys Lys Gln Arg Ala Gly Val Met Leu Thr Val Gly Thr Lys		160
	165	170	175
10	Pro Asp Gly Ser Pro Asp Glu Arg Trp Cys Phe Arg Val Glu Glu Val		
	180	185	190
	Asn Trp Ala Ser Trp Glu Gln Thr Leu Pro Thr Leu Cys Glu Asp Pro		
	195	200	205
	Ser Gly Ala Gly Val Pro Arg Thr Leu		
15	210	215	

<210> 5
 <211> 2348
 <212> DNA
 20 <213> Homo sapiens

	<400> 5	
	taggatgacc tcaccctcca gctctccagt tttcagggtt gagacattag atggaggcca	60
	agaagatggc tctgaggcg gg acagaggaaa gctggatttt gggagcgggc tgcctccat	120
25	ggagtcacag ttccagggcg aggaccggaa attcgcctt cagataagag tcaacctcaa	180
	ctaccgaaag ggaacaggtg ccagtcagcc ggatccaaac cgatttgc acc gagatcggt	240
	cttcaatgcg gtctcccg ggatctggct ggacttccag agtacctgag	300
	caagaccagc aagtaccta cccactcgga atacacagag ggctccacag gtaagacgtg	360
	cctgtatgaag gctgtgctga accttaagga cggagtcaat gcctgcattc tgccactgt	420
30	gcagatcgac aggactctg gcaatcctca gcccctggta aatgcccagt gcacagatga	480
	ctattaccga ggcacacgcg ctctgcacat cgccatttgg aagaggagtc tgcagtgtgt	540
	gaagctcctg gtggagaatg gggccaatgt gcatgcccgg gcctgcggcc gcttcttcca	600
	gaaggcccaa gggacttgc tttatccgg tgagctaccc ctctcttgg ccgcttgcac	660
35	caagcagtgg gatgtggtaa gctaccttctt ggagaaccca caccagcccg ccagcctgca	720
	ggccactgac tcccaggcgca acacagtctt gcatgcccata gtgatgatct cggacaactc	780
	agctgagaac attgcactgg tgaccagcat gatgtatggg ctcttccaag ctggggcccg	840
	cctctgcctt accgtgcagc ttgaggacat ccgcaacctg caggatctca cgcctctgaa	900
	gctggccgccc aaggaggcgca agatcgagat tttcaggcac atcctgcagc gggagtttgc	960
40	aggactgagc caccttcccc gaaagttcac cgagtggtgc tatggccctg tccgggtgtc	1020
	gctgtatgac ctggcttctg tggacagctg tgaggagaac tcagtgctgg agatcattgc	1080
	ctttcattgc aagagccccgc accgacacccg aatggtcgtt ttggagcccc tgaacaaact	1140

gctgcaggcg	aatgggatc	tgctcatccc	caagtttttc	ttaaacttcc	tgtgtaatct	1200	
gatctacatg	ttcatcttca	ccgctgttgc	ctaccatca	cctaccctga	agaaggccgc	1260	
ccctcacctg	aaagcgagg	ttggaaactc	catgctgtg	acggggccaca	tccttacct	1320	
gctagggggg	atctacctcc	tcgtgggcca	gctgtggtac	ttctggcggc	gccacgtgtt	1380	
5	catctggatc	tcgttcatag	acagctactt	tgaaatccctc	ttcctgttcc	aggccctgt	1440
	cacagtggtg	tcccaggtgc	tgtgtttcct	ggccatcgag	tggtacctgc	ccctgcttgt	1500
	gtctgcgtg	gtgtgggct	ggctgaacct	gtttactat	acacgtggct	tccagcacac	1560
	aggcatctac	agtgtcatga	tccagaaggt	catcctgcgg	gacctgctgc	gtttccttct	1620
10	gatctactta	gtttccttt	tcggcttcgc	tgtagccctg	gtgagcctga	gccaggaggc	1680
	ttggcgcccc	gaagctccta	caggccccaa	tgccacagag	tcaagtgcagc	ccatggaggg	1740
	acaggaggac	gagggcaacg	gggcccagta	caggggtatc	ctggaaagcct	ccttgagct	1800
	cttcaattc	accatcggca	tgggcgagct	ggccttccag	gagcagctgc	acttccgcgg	1860
	catggtgctg	ctgctgctgc	tggcctacgt	gctgctcacc	tacatcctgc	tgctcaacat	1920
15	gctcatcgcc	ctcatgagcg	agaccgtcaa	cagtgcgcc	actgacagct	ggagcatctg	1980
	gaagctgcag	aaagccatct	ctgtcctgga	gatggagaat	ggctatttgt	ggtgcaggaa	2040
	gaagcagcgg	gcaggtgtga	tgctgaccgt	tggcactaag	ccagatggca	gccccgatga	2100
	gctgctggtgc	ttcaggggtgg	aggaggtgaa	ctgggcttca	tgggagcaga	cgctgcctac	2160
	gctgtgtgag	gaccgcgtcag	gggcaggtgt	ccctcgaact	ctcgagaacc	ctgtcctggc	2220
20	ttccccctccc	aaggaggatg	aggatggtgc	ctctgaggaa	aactatgtgc	ccgtccagct	2280
	cctccagtc	aactgtatggc	ccagatgcag	caggaggcca	gaggacagag	cagaggatct	2340
	ttccaacc						2348

<210> 6

<211> 763

25 <212> PRT

<213> Homo sapiens

<400> 6

Met Thr Ser Pro Ser Ser Pro Val Phe Arg Leu Glu Thr Leu Asp

30	1	5	10	15
----	---	---	----	----

Gly Gly Gln Glu Asp Gly Ser Glu Ala Asp Arg Gly Lys Leu Asp Phe

20	25	30
----	----	----

Gly Ser Gly Leu Pro Pro Met Glu Ser Gln Phe Gln Gly Glu Asp Arg

35	40	45
----	----	----

35	Lys Phe Ala Pro Gln Ile Arg Val Asn Leu Asn Tyr Arg Lys Gly Thr	
50	55	60

Gly Ala Ser Gln Pro Asp Pro Asn Arg Phe Asp Arg Asp Arg Leu Phe

65	70	75	80
----	----	----	----

40	Asn Ala Val Ser Arg Gly Val Pro Glu Asp Leu Ala Gly Leu Pro Glu	
85	90	95

Tyr Leu Ser Lys Thr Ser Lys Tyr Leu Thr Asp Ser Glu Tyr Thr Glu

	100	105	110
	Gly Ser Thr Gly Lys Thr Cys Leu Met Lys Ala Val Leu Asn Leu Lys		
	115	120	125
5	Asp Gly Val Asn Ala Cys Ile Leu Pro Leu Leu Gln Ile Asp Arg Asp		
	130	135	140
	Ser Gly Asn Pro Gln Pro Leu Val Asn Ala Gln Cys Thr Asp Asp Tyr		
	145	150	155
	Tyr Arg Gly His Ser Ala Leu His Ile Ala Ile Glu Lys Arg Ser Leu		
	165	170	175
10	Gln Cys Val Lys Leu Leu Val Glu Asn Gly Ala Asn Val His Ala Arg		
	180	185	190
	Ala Cys Gly Arg Phe Phe Gln Lys Gly Gln Gly Thr Cys Phe Tyr Phe		
	195	200	205
15	Gly Glu Leu Pro Leu Ser Leu Ala Ala Cys Thr Lys Gln Trp Asp Val		
	210	215	220
	Val Ser Tyr Leu Leu Glu Asn Pro His Gln Pro Ala Ser Leu Gln Ala		
	225	230	235
	Thr Asp Ser Gln Gly Asn Thr Val Leu His Ala Leu Val Met Ile Ser		
	245	250	255
20	Asp Asn Ser Ala Glu Asn Ile Ala Leu Val Thr Ser Met Tyr Asp Gly		
	260	265	270
	Leu Leu Gln Ala Gly Ala Arg Leu Cys Pro Thr Val Gln Leu Glu Asp		
	275	280	285
25	Ile Arg Asn Leu Gln Asp Leu Thr Pro Leu Lys Leu Ala Ala Lys Glu		
	290	295	300
	Gly Lys Ile Glu Ile Phe Arg His Ile Leu Gln Arg Glu Phe Ser Gly		
	305	310	315
	Leu Ser His Leu Ser Arg Lys Phe Thr Glu Trp Cys Tyr Gly Pro Val		
	325	330	335
30	Arg Val Ser Leu Tyr Asp Leu Ala Ser Val Asp Ser Cys Glu Glu Asn		
	340	345	350
	Ser Val Leu Glu Ile Ile Ala Phe His Cys Lys Ser Pro His Arg His		
	355	360	365
35	Arg Met Val Val Leu Glu Pro Leu Asn Lys Leu Leu Gln Ala Lys Trp		
	370	375	380
	Asp Leu Leu Ile Pro Lys Phe Phe Leu Asn Phe Leu Cys Asn Leu Ile		
	385	390	395
	Tyr Met Phe Ile Phe Thr Ala Val Ala Tyr His Gln Pro Thr Leu Lys		
	405	410	415
40	Lys Ala Ala Pro His Leu Lys Ala Glu Val Gly Asn Ser Met Leu Leu		
	420	425	430

Thr Gly His Ile Leu Ile Leu Leu Gly Gly Ile Tyr Leu Leu Val Gly
435 440 445
Gln Leu Trp Tyr Phe Trp Arg Arg His Val Phe Ile Trp Ile Ser Phe
450 455 460
5 Ile Asp Ser Tyr Phe Glu Ile Leu Phe Leu Phe Gln Ala Leu Leu Thr
465 470 475 480
Val Val Ser Gln Val Leu Cys Phe Leu Ala Ile Glu Trp Tyr Leu Pro
485 490 495
Leu Leu Val Ser Ala Leu Val Leu Gly Trp Leu Asn Leu Leu Tyr Tyr
10 500 505 510
Thr Arg Gly Phe Gln His Thr Gly Ile Tyr Ser Val Met Ile Gln Lys
515 520 525
Val Ile Leu Arg Asp Leu Leu Arg Phe Leu Leu Ile Tyr Leu Val Phe
530 535 540
15 Leu Phe Gly Phe Ala Val Ala Leu Val Ser Leu Ser Gln Glu Ala Trp
545 550 555 560
Arg Pro Glu Ala Pro Thr Gly Pro Asn Ala Thr Glu Ser Val Gln Pro
565 570 575
Met Glu Gly Gln Glu Asp Glu Gly Asn Gly Ala Gln Tyr Arg Gly Ile
20 580 585 590
Leu Glu Ala Ser Leu Glu Leu Phe Lys Phe Thr Ile Gly Met Gly Glu
595 600 605
Leu Ala Phe Gln Glu Gln Leu His Phe Arg Gly Met Val Leu Leu Leu
610 615 620
25 Leu Leu Ala Tyr Val Leu Leu Thr Tyr Ile Leu Leu Leu Asn Met Leu
625 630 635 640
Ile Ala Leu Met Ser Glu Thr Val Asn Ser Val Ala Thr Asp Ser Trp
645 650 655
Ser Ile Trp Lys Leu Gln Lys Ala Ile Ser Val Leu Glu Met Glu Asn
30 660 665 670
Gly Tyr Trp Trp Cys Arg Lys Lys Gln Arg Ala Gly Val Met Leu Thr
675 680 685
Val Gly Thr Lys Pro Asp Gly Ser Pro Asp Glu Arg Trp Cys Phe Arg
690 695 700
35 Val Glu Glu Val Asn Trp Ala Ser Trp Glu Gln Thr Leu Pro Thr Leu
705 710 715 720
Cys Glu Asp Pro Ser Gly Ala Gly Val Pro Arg Thr Leu Glu Asn Pro
725 730 735
Val Leu Ala Ser Pro Pro Lys Glu Asp Glu Asp Gly Ala Ser Glu Glu
40 740 745 750
Asn Tyr Val Pro Val Gln Leu Leu Gln Ser Asn

755

760

Claims

1. An isolated polypeptide comprising an amino acid sequence which has at least 95% identity to the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:6 over the entire length of SEQ ID NO:2 or SEQ ID NO:6, respectively.
2. An isolated polypeptide as claimed in claim 1 in which the amino acid sequence has at least 98% identity.
- 10 3. The polypeptide as claimed in claim 1 comprising the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:6.
4. The isolated polypeptide of SEQ ID NO:2 or SEQ ID NO:6.
- 15 5. An isolated polynucleotide comprising a nucleotide sequence encoding a polypeptide that has at least 95% identity to the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:6, over the entire length of SEQ ID NO:2 or SEQ ID NO:6, respectively; or a nucleotide sequence complementary to said isolated polynucleotide.
- 20 6. An isolated polynucleotide comprising a nucleotide sequence that has at least 95% identity to a nucleotide sequence encoding a polypeptide of SEQ ID NO:2 or SEQ ID NO:6, over the entire coding region; or a nucleotide sequence complementary to said isolated polynucleotide.
- 25 7. An isolated polynucleotide which comprises a nucleotide sequence which has at least 95% identity to that of SEQ ID NO:1 or SEQ ID NO:5 over the entire length of SEQ ID NO:1 or SEQ ID NO:5, respectively; or a nucleotide sequence complementary to said isolated polynucleotide.
- 30 8. The isolated polynucleotide as claimed in any one of claims 5 to 7 in which the identity is at least 98%.
9. An isolated polynucleotide selected from:
 - (a) a polynucleotide comprising a nucleotide sequence encoding the polypeptide of SEQ ID NO:2 or SEQ ID NO:6;
 - (b) the polynucleotide of SEQ ID NO:1 or SEQ ID NO:5; and

(c) a polynucleotide obtainable by screening an appropriate library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NO:1 or SEQ ID NO:5 or a fragment thereof;
or a nucleotide sequence complementary to said isolated polynucleotide

5

10. An expression system comprising a polynucleotide capable of producing a polypeptide of claim 1 when said expression system is present in a compatible host cell.

11. A host cell comprising the expression system of claim 10 or a membrane thereof expressing 10 the polypeptide of claim 1.

12. A process for producing a polypeptide of claim 1 comprising culturing a host cell of claim 11 under conditions sufficient for the production of said polypeptide and recovering the polypeptide from the culture medium.

15

13. An antibody immunospecific for the polypeptide of claim 1.

14. A method for screening to identify compounds which stimulate or which inhibit the function of the polypeptide of claim 1 which comprises a method selected from the group consisting of:

20 (a) measuring the binding of a candidate compound to the polypeptide (or to the cells or membranes bearing the polypeptide) or a fusion protein thereof by means of a label directly or indirectly associated with the candidate compound;

(b) measuring the binding of a candidate compound to the polypeptide (or to the cells or membranes bearing the polypeptide) or a fusion protein thereof in the presence of a labeled 25 competitor;

(c) testing whether the candidate compound results in a signal generated by activation or inhibition of the polypeptide, using detection systems appropriate to the cells or cell membranes bearing the polypeptide;

(d) mixing a candidate compound with a solution containing a polypeptide of claim 1, to form a 30 mixture, measuring activity of the polypeptide in the mixture, and comparing the activity of the mixture to a standard; or

(e) detecting the effect of a candidate compound on the production of mRNA encoding said polypeptide and said polypeptide in cells, using for instance, an ELISA assay.

15. A process for diagnosing a disease or a susceptibility to a disease in a subject related to expression or activity of the polypeptide of claim 1 in a subject comprising:

(a) determining the presence or absence of a mutation in the nucleotide sequence encoding said polypeptide in the genome of said subject; and/or

5 (b) analyzing for the presence or amount of said polypeptide expression in a sample derived from said subject.

16. An isolated polynucleotide selected from the group consisting of

(a) an isolated polynucleotide comprising a nucleotide sequence which has at least 95% identity 10 to SEQ ID NO:3 over the entire length of SEQ ID NO:3;

(b) an isolated polynucleotide comprising the polynucleotide of SEQ ID NO:3;

(c) the polynucleotide of SEQ ID NO:3; or

(d) an isolated polynucleotide comprising a nucleotide sequence encoding a polypeptide which has 15 at least 95% identity to the amino acid sequence of SEQ ID NO:4, over the entire length of SEQ ID NO:4.

17. A polypeptide selected from the group consisting of

(a) a polypeptide which comprises an amino acid sequence which has at least 95% identity to that 20 of SEQ ID NO:4 over the entire length of SEQ ID NO:4;

(b) a polypeptide in which the amino acid sequence has at least 95% identity to the amino acid sequence of SEQ ID NO:4 over the entire length of SEQ ID NO:4;

(c) a polypeptide which comprises the amino acid of SEQ ID NO:4;

(d) a polypeptide which is the polypeptide of SEQ ID NO:4; or

(e) a polypeptide which is encoded by a polynucleotide comprising the sequence contained in 25 SEQ ID NO:3.

INTERNATIONAL SEARCH REPORT

Int. Application No
PCT/EP 99/00420

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C12N15/11 C12N15/85 C12N5/10 C07K14/705

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CATERINA M J ET AL: "THE CAPSAICIN RECEPTOR: A HEAT-ACTIVATED ION CHANNEL IN THE PAIN PATHWAY" NATURE, vol. 389, no. 1, 23 October 1997, pages 816-824, XP002915410 cited in the application see page 816 - page 823; figure 5 ---	1-17
P, X	WO 98 39448 A (HUMAN GENOME SCIENCES, INC. ET AL) 11 September 1998 Sequence ID. 191 Sequence ID. 191 see claims 1-23 ---	16, 17
A	---	1-15 -/-

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
18 June 1999	30/06/1999
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040. Tx. 31 651 epo nl. Fax: (+31-70) 340-3016	Authorized officer van Klompenburg, W

INTERNATIONAL SEARCH REPORT

Int. Application No.
PCT/EP 99/00420

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E	WO 99 09140 A (BRAKE ANTHONY ; JULIUS DAVID J (US); UNIV CALIFORNIA (US); CATERINA) 25 February 1999 sequence id. 5-9,20-23 see claims 1-19 ----	1-17
T	CATERINA ET AL: "A capsaicin-receptor homologue with a high threshold for noxious heat" NATURE, vol. 398, 1 April 1999, pages 436-441, XP002105951 see page 436 - page 441; figures 1,5 ----	1-17

INTERNATIONAL SEARCH REPORT

International application No.

PCT/EP 99/00420

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claim 15 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. Claims Nos.: because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte [REDACTED] Application No
PCT/EP 99/00420

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9839448 A		NONE	
WO 9909140 A	25-02-1999	AU 9115698 A	08-03-1999

THIS PAGE BLANK (USPTO)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)